# An information assistant system for the prevention of tunnel vision in crisis management

Yujia Cao
Human Media Interaction (HMI)
University of Twente
7500 AE Enschede, the Netherlands
y.cao@utwente.nl

April, 2008

## Abstract

In the crisis management environment, tunnel vision is a set of bias in decision makers' cognitive process which often leads to incorrect understanding of the real crisis situation, biased perception of information, and improper decisions. The tunnel vision phenomenon is a consequence of both the challenges in the task and the natural limitation in a human being's cognitive process. An information assistant system is proposed with the purpose of preventing tunnel vision. The system serves as a platform for monitoring the on-going crisis event. All information goes through the system before arrives at the user. The system enhances the data quality, reduces the data quantity and presents the crisis information in a manner that prevents or repairs the user's cognitive overload. While working with such a system, the users (crisis managers) are expected to be more likely to stay aware of the actual situation, stay open minded to possibilities, and make proper decisions.

# Contents

# 1  Introduction

Crisis is generally understood as an urgent situation with a negative outcome. Crisis management is a strategic management activity aiming to prevent or minimize the negative impact of a crisis. It is a time-critical task with high uncertainty, high risk, and high information density (information overload). Stress and possible cognitive overload may lead to incorrect understanding of the actual crisis situation and biased confirmation of information, as we call "tunnel vision phenomena". Computers have several advantages over the human brain which have great value for the prevention of tunnel vision. For example, they are better at information recording, multi-tasking and complex calculation. They act only on logic, but not on emotion. Aiming at the prevention of tunnel vision, we propose an information assistant system to enhance the information quality, deal with information overload, and reduce the user's cognitive load. We believe, when working with such a system, human managers are more likely to keep a clear mind of the real situation and stay open-minded when making decisions.

The outline of the remainder is as follows: Section 2 describes the challenges of crisis management and the existing applications of IT and AI technology for the crisis management support. Section 3 introduces the phenomena and definition of tunnel vision. Cognitive psychology theories give further explanations of why tunnel vision occurs. Section 4 presents the overview of the proposed information assistant system. Sections 5-7 zoom into three aspects in the system design. Section 5 presents the representation strategy of a crisis scenario. The modality planning module, which is the research focus of the author, will be described with more details in section 6. The function and structure of the reasoning and mining module will also be further discussed in section 7. However, it will not be one part of the author's individual research.

# 2  Crisis management

## 2.1  Definitions of crisis

The word "crisis", or alternatively "disaster", "hazard", is generally understood as an urgent situation with negative outcome. Crises come in many forms, including: 1) Nature and man-made disasters, such as earthquakes, storms, floods, oil spills, chemical fires and explosions; 2) Transport accidents, such as car, train and air plane crash; 3) Terrorism and civil attacks, such as bombs and gun shootings;  4) Economic crises, such as inflation, company bankruptcy and stock market crashes.

Due to the diversity of the crisis domains, the definition of crisis differs in different fields. In economical field, a crisis event is defined as "a major occurrence with a potentially negative outcome affecting an organization, company, or industry, as well as its publics, products, services, or good name" [1] or "a low-probability, high impact event that threatens the viability of the organization and is characterized by ambiguity of cause, effect, and means of resolution, as well as by a belief that decisions must be made swiftly"

[2]. More generally, Smith [3] defined hazard as "a naturally occurring or human-induced process or event with the potential to create loss, i.e., a general source of danger". Fritz [4] defines disaster as "an event, concentrated in time and space, in which a society, or a relatively self-sufficient subdivision of a society, undergoes severe danger and incurs such losses to its members and physical appurtenances that the social structure is disturbed and the fulfillment of all or some of the essential functions of the society is prevented".

Regardless of different domain, crisis events often share several common characteristics: 1) high importance 2) high urgency 3) high uncertainty.


## 2.2 Crisis management

Crisis management, or in another name – emergency management, is "the discipline and profession of applying science, technology, planning and management to deal with extreme events that can injure or kill large numbers of people, do extensive damage to property, and disrupt community life" [5]. It is a strategic management activity aiming to prevent or minimize the negative impact of a crisis. The characteristics of a crisis event make the domain extremely challenging. Here we highlight a set of main challenges.

1) Time critical

Crisis management is a time-dependent task. Time is an essential issue during the whole crisis management process. It is measured in minutes and seconds. A new-forming crisis requires concern and response as soon as possible. Timely and correct decisions may shorten the crisis lifetime and reduce the negative impact. It is not exaggerated to say "time is money" and "time is life".

2) High risk

Crisis management is a high-risk activity. Decision makers know that their unsuitable decisions would cause more loss or even lives. High risk brings great stress which is a mental obstacle for the crisis manager to give a high quality performance.

3) Large quantity, low quality information

Crisis managers typically have to deal with information overload [6-9]. Decision makers face large amounts of historical information to dig through. New information is being added with a high rate. It is difficult to identify what information is relevant to the decision. To make it worse, the quality of information is often very low. They may be distorted, misrouted, or concealed [6, 9]. Inconsistent information chunks needs to be pieced together into an accurate understandable picture of reality.

4) large uncertainty

Crises are low frequency events; as a result, almost everything in a crisis is an exception to the norm. Decision makers often don't have sufficient knowledge or experience to draw upon. The next event may not resemble the last event in ways that would permit theoretical constructs to be developed [10]. Future more, there is no way to predict what information is going to be needed and who is going to need it. There is no way to predict exactly who is going to be doing what, when, why and/or how at the command and control level in a crisis environment. In another word, the exact actions and responsibilities of the individual in a large crisis response group cannot be predetermined. It also means that one cannot exactly predict what they might be interested in ahead of time [11].

Crisis management is about communication and collaboration. It requires effective communication and collaboration between various players. However, the function of crisis managers is especially important. Except the historical cases they experienced, the simulation trainings they had, the predesigned routine they learnt, their performance for the next emergence response task is also a function of the following factors:

1) The stress level of the crisis manager

When an event is unexpected or novel, lack of decision readiness induces stress. The less familiar the crisis event is and the less ready decision makers are to deal with it, the greater the stress that they experience [9]. The stress results in changes in both the information and control processes of a person, and, because of these changes, a person's behavior becomes less varied or flexible [12].

2) The quality of information crisis manager receives

The quality of information is a very important factor of the quality of the crisis decision. In order to improve of the response process, it is extremely important to have an information aid which can help the crisis manager to learn and understand what actually happened before, during, and after the crisis[11].

## 2.3  IT&AI for crisis management

Crisis management is in nature a multi-disciplinary research topic. Research efforts from different domain need to be joined effectively to fulfill the comprehensive task. The use of information and communication technology has been applied to all phases of crisis management: preparedness, planning, training, response, recovery, and assessment [8].

A set of IT-based crisis planning tools has been described by Nunamaker [9] for organizations to become more intelligent by supporting and improving their cognitive activities. The tools operate a sequence of carefully designed cognitive events to help decision groups acquire expertise in managing crises and learn how to deal with any situation. Desimone [13] discussed the application of artificial intelligence planning technology to a crisis management task in the context of several cases studies involving military and non-military crisis management activates.

In recent years, Agent technology [14] has been used to support many processes throughout the crisis management cycle from mitigation and preparation to actual response and recovery [15]. Main research areas are: 1) agent-based simulation systems [16, 17] 2) agent-based decision support systems [18] 3) agent-based network-centric systems [19-22].

Decision making support is another topic which has gained much research effort. Chapman [23] described a decision making aid by providing focused mining of unstructured data to create a timely structured data repository. Semantic Web and natural language technologies are applied. Shilov [18] presented a decision mining method which finds the common preferences for each role of decision making participants and automatically selects the best decisions in a typical situation based on the preferences. Data mining and context modeling technologies are applied.

Although computers play an important role in supporting emergency response, the traditional human-computer interaction (mouse and keyboard) present a bottleneck at interface level by encumbering the information exchange [24]. Hence, multimodal

human-computer interfaces are designed to allow more effective and efficient communication between a wide range of users and devices [24-27]. In these multimodal systems, current data and events are visualized in a dynamic interactive map with high-resolution. Users can naturally interact with the large screen display through gesture, speech and facial expression.

Turoff, M. et al. [11] defined a framework for the sensible development of flexible and dynamic emergency response information systems. First, they organized a set of premises and concepts which cut across all types of crisis and emergency management task. Based upon them, they systematically developed a set of general and supporting design principles and specifications.

At last, it is commonly agreed that, despite of how powerful and helpful information technology is, in emergency operations, humans are in charge of technology but not replaced by them.

# 3 Tunnel vision

## 3.1 Tunnel vision in various fields

The term "tunnel vision" originally comes from the medical domain. It was coined by Donders in 1855 and refers to a group of retinal diseases with common attributes, such as retinitis pigmentosa or glaucoma [1]. Tunnel vision patients suffer from the loss of peripheral vision with retention of central vision; therefore they have a constricted circular tunnel-like field of vision. So far the cure for tunnel vision diseases remains an unsolved problem, however several eye research institutes have been working on special glasses which open up a whole new world for tunnel vision patients[2].

The term "tunnel vision" has been adopted by many other research works outside its original domain and redefined according to different research tasks. In the website visibility research of Hasan et al [28], tunnel vision is understood as the phenomenon that website users get too familiar with the content and layout of frequently visited Web sites. It is one of the main visibility concerns facing information delivery and knowledge exchange through Web sites. Several technologies have been proposed to overcome tunnel vision and enhance the visibility of information.

Norman [29] used the term "tunnel vision" in his research on how affect influences human behavior. Negative affect focuses the human mind, leading to better concentration. It causes tunnel vision when creative problem solving is required. Norman gives an objective view of tunnel vision. He points out that it can lead to harm sometimes, but it is good when we need to concentrate attention to avoid distraction by irrelevant, extraneous matters.

Tunnel vision is also adopted by criminal justice research in the law domain. In the research works from Dianne [30] and Keith et al. [31], tunnel vision is viewed as a natural human tendency that has particularly pernicious effects in the criminal justice system. Compendiums of common heuristics and logical fallacies lead actors in the

---

[1] http://www.lowvision.org/retinitis_pigmentosa.htm
[2] http://www.wyff4.com/health/10570682/detail.html

6

criminal justice system to focus on a suspect, select and filter the evidence that will build a case for conviction, while ignoring or suppressing evidence that points away from guilt.

As a summary of above, the term "tunnel" always indicates "limited" or "biased", while "vision" evolves into "attention" or "cognition".

## 3.2 Tunnel vision in crisis management

In this research, "tunnel vision" is transplanted into the crisis management area. As discussed in the previous chapter, the characteristics of crisis events make the emergency response task very comprehensive and challenging. It is especially true to the decision making group (crisis managers) in the crisis response center. Too much information and too little time may lead to cognitive overload. In a stressful and cognitive overload situation, the decision makers often have the tendency to have biases in their situation understanding and problem solving processes. The phenomena of tunnel vision, though not the exact term, are mentioned by many emergency response related researches with different focus according to different application and domain background. We summarize the tunnel vision phenomena into two main aspects.

1) Empiricism and bias on option perception

Under stress due to the time urgency and the lack of necessary information or coordination awareness, crisis managers tend to create one coherent description of the real world without full awareness of what was actually happening. They also tend to rely on standard operating procedures, previous ways of understanding without reexamination. However almost everything in a crisis is an exception to the norm. With luck, their behavior may fit the situation. More often, though, it does not [9, 11].

2) Bias on information confirmation and option perception

Once an understanding is forming, the crisis managers tend to perceive and confirm only clear and familiar information which aligns with their subjective understanding and explanation of the crisis situation. Research has shown that correct but ambiguous or contradicting information is easily discarded, while confirmative and neutral information are used to strengthen the initial beliefs [32-34]. Once a solution is forming, they tend to stick to it without examining other possibilities. This may lead to a growing bias in the diagnosis of the real crisis situation, which in turn cause costly delay and loss [9, 11, 35].

In order to better understand the cause of tunnel vision and further explain the phenomena, literature studies into the psychology domain have been carried out. A couple of cognitive psychology theories show that human beings in nature have difficulties in reasoning, decision making and problem solving, when they are facing large amount of information, complex task, high uncertainty, stress and threat.

1) Effect of uncertainty – judgment limitation

Humans are limited to judge and make choice in complex situations. Faced with extreme uncertainty, humans tend to increase their search for information, often in self-fulfilling ways, while simultaneously shutting down some channels of communication, and relying on familiar or formal information and channels [36]. When people assess the probability of an uncertain event or the value of an uncertain quantity, they rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these

heuristics are helpful, but sometimes, these lead to severe and systematic errors [37]. In the specific task of crisis response, decision makers often construct overly simple representations of the actual problem and solve problems by breaking them up and solving the components separately and then using the aggregation as the solution. Such oversimplification does not guarantee optimal solutions to complex problems.

2) Effect of information overload - cognitive load capacity limitation

Cognitive Load refers to the load on working memory during problem solving, thinking and reasoning (including perception, memory, language, etc.). It is mainly used in educational psychology researches related to the human learning process and instruction design [38-41]. Cognitive Load Theory [38] states that optimum learning occurs in humans when the load on working memory is kept to a minimum to best facilitate the changes in long term memory. Working memory (short-term memory), which has been described as "hub of cognition", refers to the system responsible for the temporary storage and concurrent processing of information [42]. It is severely limited and serves as an important bottleneck in human information processing [43]. Therefore, any complex problem that requires a large number of information items and variables to be held in short-term memory may contribute to an excessive cognitive load. Cognitive overload occurs when available cognitive capacity is not sufficient to meet the required processing demands.

Fowlkes et. al. [44] described a cognitive load model with 3 measurement factors: 1) percentage time occupied 2) level of information processing 3) number of task-set switches. According to their model, cognitive overload can occur when the operator does not have enough time to finish the tasks, the operator tasks are too complicated or the operator has to perform too many tasks at the same time (or a combination of any of these elements).

3) Effect of stress and threat

Due to the limited cognitive capabilities, there is a general tendency for individuals, groups, and organizations to behave rigidly in stressful and threatening situations. A threat may result in restriction of information processing, such as a narrowing in the field of attention, a simplification in information codes, or a reduction in the number of channels used. Thus, human cognitive process becomes less varied or flexible [12].

The above theories provide much deeper insight on why the tunnel vision phenomenon often occurs in the crisis management environment. The phenomenon itself can also be further explained by two psychology terms, the framing bias and the confirmation bias.

1) Framing bias

A "decision frame" is the decision maker's cognitive representation of a task [45, 46]. The decision maker uses this frame to evaluate the outcomes and contingencies associated with a particular choice [47]. For example, a clinical practice frame of a physician is his mental interpretation of the results of clinical trials (numeric data). It influences the physician's perceptions on the worth of treatments. After a consumer watched a product advertisement, his mental interpretation of this advertisement (frame) influences on whether he is willing to buy the product or not. In the crisis management domain, the decision frame of a crisis manager is his mental understanding of the crisis situation. It influences his decisions on how to respond to the crisis.

"Framing" is the cognitive process of building a decision frame from the information related to a given decision problem. Framing often can be done in more than one way [47]. Any single frame yields a partial view of a problem and may evoke different choices [48]. Therefore, biases in the framing process may lead to improper decisions. The framing bias in crisis management refers to the crisis managers' tendency to create one coherent understanding of the crisis situation based on their experiences without reexamining.

2) Confirmation bias

Confirmation bias is perhaps the best known and most widely accepted notion of inferential error on human reasoning. It refers to a fundamental tendency to seek and interpret formation consistent with their current beliefs, theories or hypotheses and to avoid information and interpretations which contradict their current beliefs. This is an unconscious selection and neglect process. There springs up, also, an unconscious pressing of the theory to make it fit the facts, and a pressing of the facts to make them fit the theory [9, 49, 50]. The confirmation bias in crisis management refers to the crisis managers' tendency to confirm only the information which falls in harmony with their understandings and ignore the other information.

We define "tunnel vision" in crisis management context as "the phenomena of framing bias and confirmation bias in the crisis managers' cognitive processes due to high risk, high stress, high uncertainty and high information density caused by the crisis event". If an improper decision frame is continuously confirmed, the growing bias may lead to costly delay and errors.


## 3.3 Why and how the computer can help

Considering the specific task of crisis management, computer systems have several advantages over the human brain which have high value for the prevention of tunnel vision. First of all, the computer beats the brain with information recording. The computer is able to continuously record data into its memory with high speed, regardless of the quantity and the quality of the data. If necessary, data can be copied to external memory for long term storage. From the capacity point of view, the human brain is absolutely not a loser. The average brain can hold about 100 million megabytes information. But a task such as remembering a large chunk of incoherent information normally takes much time and effort. Without regular repeating, the memory gradually gets lost.

Second, it is impossible for the brain to act without emotions while computers act only on logic. Researchers already proved that brains act on emotions, and many of our actions are based on our emotional side. We know that computers act completely on logical bits. All actions performed by computers are based on the instructions in their coding and other factors have no effect on them. It is not always a good feature. However, in the crisis management environment, the performance of a computer will not be influenced by stress or threat.

Third, the computer is better at multi-tasking. Since the number of variables which the brain can hold in its short-term memory is limited, humans have difficulty to handle multiple tasks simultaneously. More concentration normally leads to better performance.

However, as long as the CPU power allows, the computer can process multiple tasks simultaneously without "distracting" each other.

Furthermore, the computer exceeds the brain in fulfilling complex calculation and rule-based tasks. For example, the computer can easily become an excellent chess player by loading a software program with all the rules, strategies and possibilities; while human beings need many years' training to reach the same level.

Despite all the advantages, computers can never replace human beings altogether. They fail in the comparison of learning skills and adaptability. The human brain is an absolute winner at non-assisted learning, acquiring new skills, inducing conclusions from past experience and creating new methods to deal with new situations. Computers also encounter great difficulties on some tasks which seem to be natural for us, such as natural language processing, speech and vision recognition.

As a conclusion, the goal of this research is not to develop an expert system program which replaces crisis managers to make decisions. However, the computer system works as an assistant who monitors crisis event, gathers and reasons the information, makes a relevant selection and presents to the human expert, aiming at helping him to stay aware of what actually happened and stay open-minded.

## 4   The information assistant system

When facing a task with high time urgency, high risk and high information density, such as crisis management, human beings show a set of cognitive biases in their problem solving processes, as we called "tunnel vision". Since this research intends to repair and prevent tunnel vision by providing a multimodal information assistant service.

As crisis events are diverse, the information involved in different types of crisis may have very different features. Before designing an information assistant system, the first thing to determine is the type of information the system will be dealing with. This research is a part of ICIS project[3], the currently-used ICIS common crisis scenario is adopted for this research. Briefly, it is about a fire turning out in a tunnel, causing traffic jams and human injuries. Under the command of the crisis managers, fire team, policemen, and doctors work as a team. They successfully put out the fire, rescue all victims and recover the traffic (see appendix B for more details). The crisis information to be presented by the system includes phone calls to the response center, phone calls between any two devices on scene, video segments taken by cameras, and signals from all kinds of sensors (fire, smoke, temperature etc). The system users are crisis managers located in the crisis response center. The center control room is equipped with powerful computers and large displays. Except common communication channels such as keyboard and mouse, the user can also interact with the system via speech and gesture.

Figure 1 shows the structure of the proposed information assistant system (also referred as "the system" in the remainder). It can be viewed as a platform for monitoring the on-going crisis event. All data goes through the system before it arrives at the user.

---

The system enhances the data quality, reduces the data quantity and presents the crisis information in a manner that prevents or repairs the user's cognitive overload. The functions of each module are described in the following.

### 1) the multimodal data fusion module

The basic function of the multimodal data fusion module is to effectively interpret the combined semantics of multimodal inputs. Possible noises and errors contained in a single input modality can be reduced through the fusion of all input modalities [26]. As mentioned above, the inputs to be fused are speech, video fragments, and sensor signals. The human speech is interpreted through the fusion of speech signals, lip movements and gestures. The output of this module is a series of actions translated from the real-world crisis event and represented according to a pre-defined world model/ontology (section 5).

### 2) the reasoning and mining module

The reasoning and mining module plays important role in the system. First, it intends to deal with information overload by grouping reduplicate information together and provides statistics for contradicting facts. Second, it reasons about the available facts and causal relations, attempting to select information which needs more concern, and in turn, has higher presentation priority. As the effects of the foregoing two functions and the data fusing function, the crisis information quality is improved and the quantity is reduced. Furthermore, when the user consults the system via the query interface (see figure 1), this module fulfils searching processes for relevant information.

### 3) the system memory and the information query interface

We construct a crisis world model (ontology) to describe all real-world entities involved in the crisis and all possible types of actions. Based on the ontology, the system records the crisis scenario into a context-aware graph structure. The graph representation extends with time as the crisis event develops. Human beings have a limited amount of working memory, especially in a stressful situation. The memory of the system serves as an extension of the crisis managers' working memory. Via the information query interface, the users have access to all previous data. Necessary recalls can be made at any time to support their decision making.

### 4) the multimodal presentation generation module

The function of the multimodal presentation generation module is to present the crisis information in an effective and efficient manner. Effectiveness means that the information is conveyed correctly; while efficiency means that the presentation manner helps to prevent or repair the user's cognitive overload. The optimum presentation manner is achieved by two main components: the order planning module and the modality planning module. The order planning module determines whether the presentation should follow the time sequences, the causal relations, the story lines, or a mixture of those three. The goal is to assist the development of a coherent understanding. The modality planning module calculates the optimum utilization of the available modalities. Different modality planning strategies are applied to different types of action input. Each strategy contains three modes, the choice among which is based on the user's preferences and cognitive state. The default mode and its adaption to the user's

preferences (if available) are applied when no cognitive overload of the user is recognized. If the user's response gets slower and the number of pending reports/requests gets larger, the modality strategies switch to the light mode and let the most urgent and important information stand out from the others. Then the user has larger chance to recover from his cognitive overload soon.

In summary, the information output by the system has higher quality and lower quantity than the real world information input. The presentation manner helps the user to prevent or repair cognitive overload. Therefore, we believe that the proposed information assistant system makes it more likely that the crisis managers stay aware of the actual situation, stay open minded to possibilities, and make proper decisions.
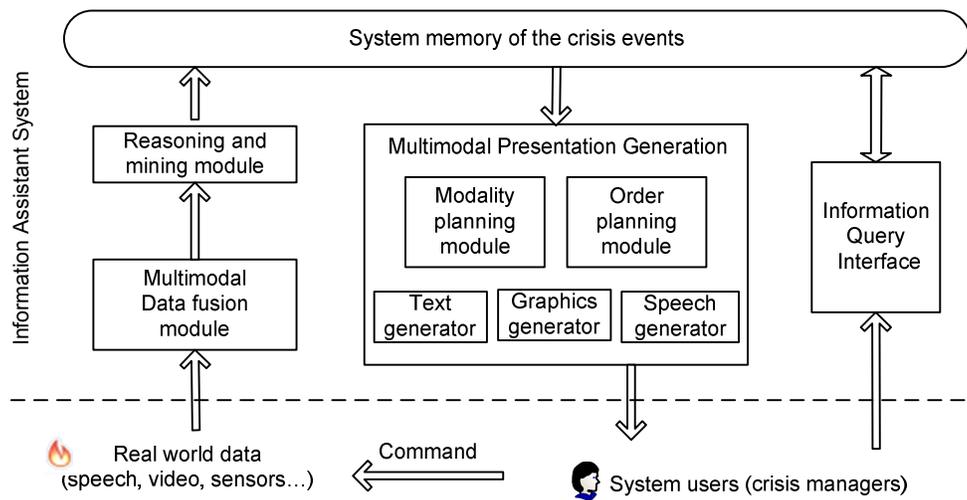


**Figure 1: Information Assistant System structure**

In the ICIS project, another multimodal framework for crisis management has been proposed with the function of facilitating the communication among all human actors and devices involved in the crisis event [26]. This framework is able to fuse multimodal inputs from a wide range of users and communication devices, and generate synchronized multimodal response for each user. However, the proposal of the information assistant system was motivated by the prevention of tunnel vision and focuses on providing information service for the crisis managers located in the crisis response center. Due to the different motivations and functions, the existing framework does not meet the major requirements of our system. These include:

- The framework does not apply any reasoning on the input data. The causal/spatial relations among the input data are not explored.
- The framework does not have a memory for recording the development of the crisis event. Only the current world state is being presented.
- The goal of the information presentation module is to deliver the same information contents and provide the same services to different communication devices. It does not address the effectiveness and the efficiency of the

12

presentation. The order planning issue and the modality planning issue are not investigated.

Therefore, the information assistant system is designed as a stand-alone system which is operated on the powerful computer in the crisis response center. We adopt the technology overlaps between the two systems, such as the multimodal data fusion technology (as described in [26]). Research efforts still need to be devoted to the following three aspects:

- Knowledge representation of the crisis scenario
- Context reasoning and graph-based data mining
- Multimodal presentation generation (modality planning and order planning)

However, it is desired that the two systems can be eventually integrated into one and provide more powerful support for the crisis managers in the crisis response center.

The multimodal presentation generation module is chosen as the research focus of the author. The design proposal of this module is introduced in section 6. Section 5 presents the knowledge representation strategy, which determines the format of the information to be presented. The function of the context reasoning and mining module will also be discussed with more details in section 7. However, it is not going to be part of the author's individual research.

# 5  Ontology-based Knowledge representation

Knowledge representation aims to create a common understanding of the real world crisis scenario between human and computer. The crisis scenario needs to be represented in a manner that a computer program can process and interpret. In this research, we use an ontology to construct the world model of the crisis scene. The model contains the description of all real world entities possibly involved in the crisis event and all possible activities. Based on the ontology, the crisis scenario is represented and recorded as two parts: 1) a world state database which contains the instances of all real world entities, and 2) a context-aware graph structure which contains all actions and the spatial/causal relations among them. This representation manner is expected to not only provide an effective memorization of the crisis development, but also enable the information mining and context reasoning for the reasoning and mining module.

## 5.1  The world model – ontology

Ontology can be generally understood as the theory of what exists. More specifically, ontology is defined as an "explicit specification of a conceptualization," which is, in turn, "the objects, concepts, and other entities that are presumed to exist in some area of interest and the relationships that hold among them." [51] In the context of artificial intelligence, ontology is referred as the shared understanding of certain domain, which is often conceived as a set of entities, relations, functions, axioms and instances. Using an ontology, we are able to [52]: (1) share common understanding of the structure of information among people or software agents, (2) enable reuse of domain knowledge (3) make domain assumptions explicit, (4) separate domain knowledge from the operational knowledge, and (5) analyze domain knowledge.

13

Ontology design is a creative process which is deeply affected by the designer's view of the domain and the potential applications. There is no single correct ontology for any domain. Fitrianie et al. [53] have designed an ontology for the crisis management domain. They employed the upper level ontology of the SmartKom [54] and a top-down development process. Physical objects in the world are categorized into 5 classes: area-perspective, object, user, communication, and history. Based on the assumption that the system flow is controlled by the dialog manager, human communications are defined as 7 dialog acts: report, prompt, answer, clarification, verification, command, and agreement. All possible processes that may be initiated by dialog manager are divided into: general process, mental process, physical process and social process. This ontology serves the task of facilitating crisis communication between human actors. It does not describe the time property of an action; therefore it is not suitable for memorizing the crisis development. It also does not describe the casual relations among actions, which is an important requirement of our ontology. As ontology design highly depends on the reference application, we need to construct an ontology for this research which serves the task of representing a crisis scenario. We use the ontology in [53] as a reference, reuse the relevant concepts, and add in our own needs.

Our ontology construction process also follows a top-down development order, starting with the definition of the most general concepts and subsequently specializing them. A concept is expressed as a class. Classes and their sub-classes are organized in a hierarchy. The top level classes are *entity* and *action*. *Entity* refers to all physical objects which involve in the crisis event, including *static entity* and *dynamic entity*. *Static entity* describes the static surroundings of crisis location, such as building, street, tunnel etc. Their properties, such as location and size, have constant value. *Dynamic entity* includes any objects which can perform an action (e.g. human actors, sensors) and/or have time varying properties (e.g. vehicles, fires). *Action* stands for all the activities carried out by human actors, such as report, request, command, rescue etc. Figure 2 shows the ontology class hierarchy.

The state of the world is formed by all entities and their properties. The properties of the dynamic entity are not constant; therefore the world state is developing. We assume that the world remains in the same state unless an action happens and changes the property of an entity. An action doesn't necessarily have influence on the state of the world (for instance a request action); however only actions can lead to state changes. In order to memorize the development of the world state and the causal relations of major changes, we define '*state_change*' as an additional concept, with the properties '*entity*', '*property*', '*new_value*', '*time*', and '*caused_by*'. The initial world state (initial property values of all entities) is stored by the system. Given any time stamp in the past, the world state can be reconstructed by the initial world state and the *state_change* instances taking place before this time stamp.

The feature of a concept and the relations between concepts are described by properties. Each concept owns certain properties and passes them down to its sub-concept. Except the inherited properties, sub-concepts may also have other own properties. The value of a property can be a data type or an instance of other classes. As shown in Figure 3, the value of property '*actor*' in class *action* is a dynamic entity instance; the value of the '*caused_by*' property can be either an entity or another action. For instance, a fire (entity) causes a sensor alarm (action). The sensor alarm (action), in turn, causes a

command from the crisis manager (action). More detailed ontology class hierarchy and properties are shown in Appendix A.
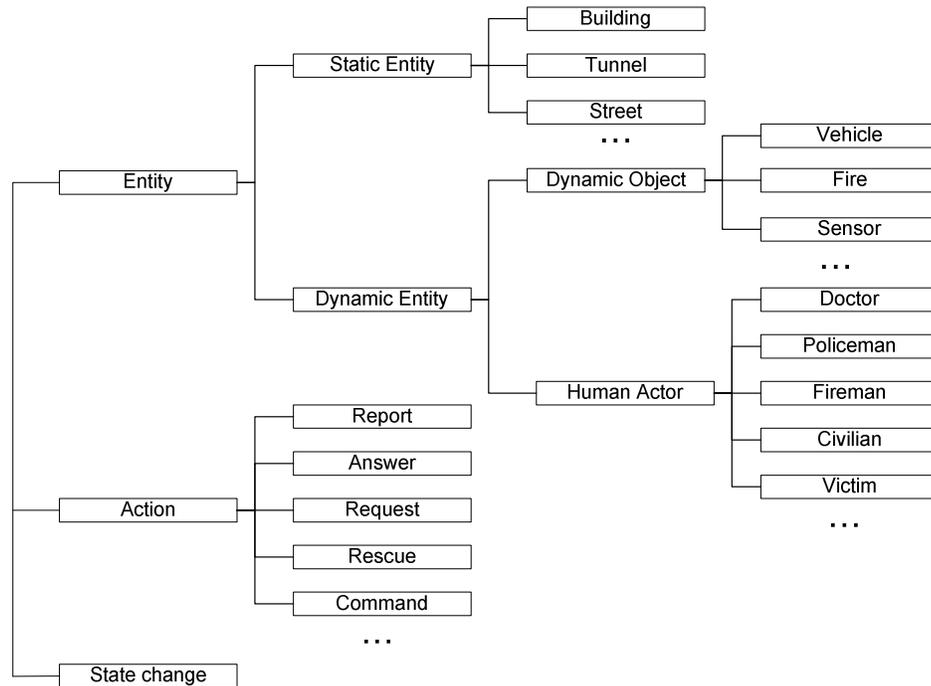


**Figure 2: Ontology class hierarchy**
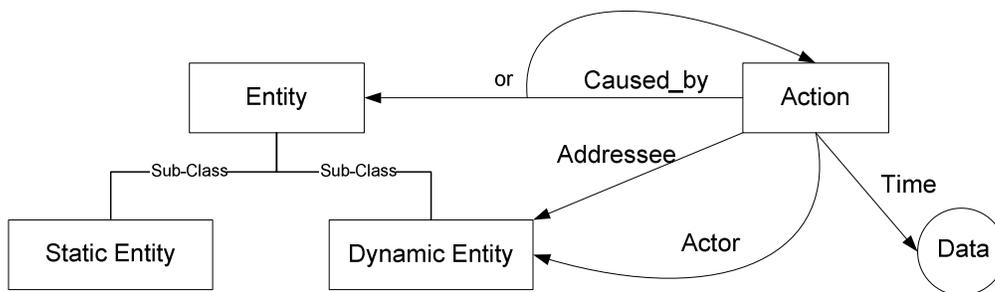


**Figure 3: Property can be data type and/or instance**

## 5.2 Context-aware representation

Context-aware means "to use context to provide relevant information and/or services to the user, where relevancy depends on the user's task [55]." In this research, context can be understood as the correlations attached to every piece of information, including the time sequence and the causal relations. The advantage of context-aware representation is that it brings the opportunities to further apply context reasoning and data mining on the represented knowledge [56].

Graph representation is widely used in the context-aware computing domain. Graphs mostly contain a set of nodes and oriented arcs. Gu [57] presented a graph representation of their context model based on a home scenario. Nodes refer to individuals, while arcs indicate properties and relations (such as defined, deduced, and sensed). Li [58] defined a context map for modeling scenes of the real world. Nodes represent activity, place, people and object, while edges indicate relations and the valid period of a relation. Contextual graphs have been applied to several operational processes, such as a coffee preparation procedure [59] and security policy [60]. A contextual graph contains a parallel organization of nodes connected by oriented arcs. It allows a context-based representation of a given problem solving for operational processes by taking into account the working environment.
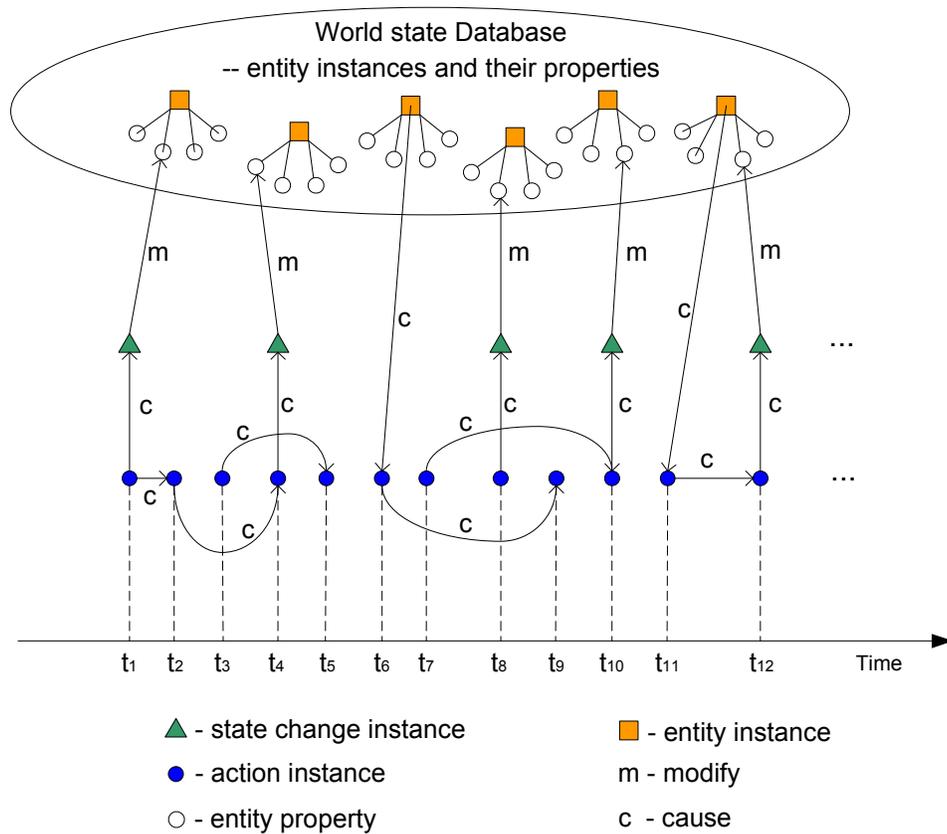


**Figure 4: Knowledge representation of the real world crisis scenario**

In this research, we also use a context-aware graph structure, together with a world state database, to represent a crisis scenario (see figure 4). The graph structure records the real world actions (nodes) and the spatial/causal relations among them (arcs). The world state database contains all entity instances and their property values. If an action has a certain influence on the state of the world, a *state_change* instance is generated, which in turn updates the properties of the related instance in the world state database. As the crisis event develops, the graph structure continuously extends and updates the world state

database. Any previous action can be directly retrieved from the graph structure. The world state at any given time stamp in the past can also be reconstructed from the initial world state (stored by the system) and the *state_change* instances taking place before this time stamp. In this way, the on-going crisis event can be represented and recorded into the system memory. Via an information query interface, the system users are able to review any previous fact. This knowledge representation manner also determines the input format of the multimodal presentation generation module. The presentation task (information to be presented) will be a series of crisis actions.

# 6  Multimodal presentation generation

The function of the multimodal presentation generation module is to present the crisis information in an effective and efficient manner. Effectiveness means that the information is conveyed correctly; while efficiency means that the presentation manner helps to prevent or repair the user's cognitive overload. The optimum presentation manner is achieved by two main components: the order planning module and the modality planning module. They are described in the following two sub-sections, respectively. Due to the current research progress, the discussion on the modality planning module is much more elaborated.

## 6.1  Modality planning

A modality is a way of presenting information that depends not only on the physical representation (the medium) but also on how it is processed (by human or machine) [61]. According to the five human perception senses (vision, hearing, touch, smell and taste), the output modalities of computer systems can be visual, auditory, tactile, olfactory and gustatory. HCI researches have shown that the usage of multiple modalities brings advantages from practical, biological and mathematical aspects [62]. Multimodal user interfaces evidently enhance ease of use, transparency, flexibility and efficiency [63]. They also permit sufficient flexibility for users to avoid many errors [64] and reduce the cognitive memory load through sharing the load of different perceptional channels [65].

Once more than one modality is available for a computer system to generate its output, the question that naturally follows is that: how can these modalities be used to best perform a certain presentation task? The problem of modality planning can be generally described as (based on [66]):

*Given a particular user, task, application domain, and/or any other aspect of computer-user interaction that is considered critical for a particular application, determine the optimum allocation and combination of modalities, to convey given information that has certain characteristics.*

Modality planning is generally agreed to be highly complex, due to many issues involved. Some factors which need to be taken into consideration are [61, 65, 67-70]:

1) The characteristics of the information to be conveyed
2) The goal of the presentation
3) The characteristics of each available modality

4) The perceiver's interests and abilities (user profile)

5) The user's cognitive load status

The modality planning task is often divided into two phases: modality allocation and modality combination. Descriptions of these two stages differ in the existing literatures. In this research, we adopt the description from Elouazizi et al. [65], which best aligns with our presentation goal.

> *Modality allocation is the process that assigns for each mode of a presentation structure the modality or combination of modalities that can best encode the content of information that the presentation has to convey. Modality combination is the process that identifies for each particular node of a presentation structure whether the modality allocation generated is a cognitively overloading combination and if so reduces the number of presentation forming modalities to only those which can complement each other for most effective processing of the information by the user.*

Take a closer look at these two phases. The modality allocation phase focuses on the "meaning-making" aspect; while the modality combination phase focuses on the cognitive load aspect. This leads exactly to our presentation desirability – effectiveness and efficiency.


### 6.1.1  Related work

The modality planning task is generally agreed to be highly complex, due to many issues involved. So far, no solution in the form of a generally applicable automated modality planner has been devised yet. Instead, most of the existing approaches focus on a small set of modalities and a certain type of application. The modality planning task is commonly considered as a mapping process from the presentation task (convey certain information) domain to the modality domain, based on pre-designed rules or strategies.

Robert [71] proposed a 3-step method to select modalities, dealing with modality expressiveness, modality effectiveness and modality adaptation, respectively. In the first step, each modality selects a range of the information it can express. The rules of expressiveness proposed in [72] are applied. The second step adjusts the amount of redundancy in order to form an effective presentation. Certain types of redundancy are considered as necessary, e.g. the additional indications for the relations between information presented in different modalities. The redundancy check procedure was controlled by a set of rules, such as a) given graphical and textual redundancy favor graphical modality, b) allow for pictorial and textual redundancy to connect the two modalities, for example. The third step adapts the final selection result to the user preference or experience.

Kerpedijev et al. [69] proposed a framework for automatic presentation of large data sets, using nature language and information graphics. Modalities are allocated by rules, witch consider the presentation effectiveness as a function of parameters such as level of detail, cardinality of the datasets being presented, and accuracy. They consider that graphics provides more compact and rapidly searchable presentations; while language is superior to graphics when the communicative function must be explicitly conveyed to the user. For different tasks, the allocation rules are designed differently. Concerning the

cardinality of the dataset being presented, allocation rules were designed quantitatively for each task.

Arens et al. [67] argue that media allocation rules should take into account not only the characteristics of the data, but also a more global view of the structure of the discourse. They first defined a discourse structure which embeds data and the presenter's goals using purpose/relations. It provides the basic organization of the information to be presented and is neutral with respect to the modality and presentation layout. Following, a presentation structure was also defined to indicate the presentation details, e.g. modality, position, content. Finally, the presentation planner transforms a discourse structure into a presentation structure by applying approximately a dozen modality allocation rules. The procedure traverses the discourse structure bottom-up, creating presentation structure nodes at the leaves of the discourse structure and then gradually filtering them up-ward until finally the entire presentation structure emerges at the root mode. Examples showed that the presentation results were better compared to direct data-to-modality mappings.

The COMET [73] system plans instructions for operating military radio, using text and graphics. After informal experiments and literature survey, six types of information units were distinguished in the representation structure. Modality allocation rules were defined for each type separately. For example, graphics alone for location and physical attributes, and text alone for communicating abstract actions and expressing connectives that indicate relationships among actions. Both text and graphics represent simple and compound actions. After modality allocation, the media coordinator combined the use of different modalities in a single explanation. By sharing a common content description, text modality and graphics modality interacted with each other bidirectionally and coordinated the sentence breaks with picture breaks and cross-references.

In the WIP system [74], a set of presentation strategies has been defined for all presentation tasks. They are represented by a name, a header, the applicability conditions and a specification of modality choice. When the presentation planner receives a presentation task, it tries to match a presentation strategy which has the corresponding effect or header. When there are more than one matches, pre-defined meta-rules are applied to make a choice.

In the SmartKOM system [75], based on 121 presentation strategies, the presentation planner recursively decomposes a high-level presentation task into primitive presentation tasks and allocates different output modalities to each primitive presentation task.

In the EMBASSI system [76], the combination of several unimodalities is defined as a multimodality. The model of a multimodality includes the set of unimodalities, the combination strategy, and the assignment to a physical output device. The combination strategy describes the synchronization, the necessary coordinations for multimodal references to objects, and the possible cross-modal references of the unimodalities. When receiving a presentation task, the presentation planner examines the user preferences and the output device condition, and then assigns one or more multimodalities, and constructs the presentation according to the combination strategies.

Zhou et al. proposed a graph-matching approach to dynamically create multimodal response to support context-sensitive information seeking in large and complex data sets [77]. Data content to be conveyed and available modalities are represented by two graphs. Data graph consists of data items and relationships between them. The media graph contains available modalities and the similarity and compatibility between two media. In

order to model all constraints, a set of metrics have been defined to assess the desirability of modality planning, including two selecting metrics and several coordination metrics. The first selection metric assesses one-to-one mapping. It takes into account factors like task-media, user-media and data-media compatibility. The second selection metric evaluates cross-media mappings. It considers the presentation recallability and affordance. Coordination metrics have been defined by data relationships, such as data importance order, data dependency, and data similarity. Finally, they apply a probabilistic graph-matching algorithm (graduated assignment algorithm [78]) to find a set of data-media mapping that maximizes the overall desirability. The algorithm iteratively applies the gradient descent method to increase the mapping probability in each step until it converges and outputs the presentation plan which maximizes the overall objective function.

According to Bernsen [79], "the best current modality planning approach" is to analyze and publicize "good compounds". To obtain the "good compounds", it is necessary to build systematic overviews of modality combinations, which have proved useful for a broad range of specified purposes. "*At least, when using a modality combination which has been certified as a good one under particular circumstances, developers will know that they are not venturing into completely unexplored territory but can make the best use of what is already known about their chosen modality combination*". For example, the combination of linguistic modality and analogue modality is one of the most commonly used "good compounds" [73-76]. Linguistic modalities (e.g. text, discourse) surpass analogue modalities (e.g. images, graphics, diagrams) at explaining abstract concepts; while analogue modalities are better at expressing what things exactly look like. Their combination may have superior expressive power [80]. Another example of a "good compound" is the combination of visual modality and auditory modality. The dual coding theory of Paivio [81] claims that humans possess separate information processing channels for visual and auditory material. Therefore, working memory has partially independent processors for handling visual and auditory signals. Mousavi et al. [82] suggest that the mixed use of both modalities can reduce cognitive load, because more effective cognitive capacity is available. For this research, a more elaborated survey on the existing modality usage guidelines is desired.


### 6.1.2 Modality taxonomy

Provoked by the thought that the modality planning task should be addressed in a unified and systematized manner, Bernsen [83] proposed a taxonomy of unimodal output modalities which serves as a theoretical foundation for understanding and generating multimodal output. Based on the observation that different modalities have different representational power, Bernsen defined the following set of basic representational properties to identify modalities.

- Linguistic / non-linguistic
- Analogue / non-analogue
- Arbitrary / non-arbitrary
- Static / dynamic
- Visual / auditory / tactile

It is claimed that this modality classification is complete, unique, relevant, and intuitive [79]. Based on this set of properties, totally 48 combinations can be derived. After some closer analysis, several removals and fusions are made, which reduce the number of combinations from 48 to 20. The 20 categories at the generic level are organized into 4 categories at the super level, as shown in table 1. The taxonomy further derived into 46 categories at the atomic level. Possible extensions to a subatomic level have also been suggested.

| SUPER LEVEL | GENERIC UNIMODAL LEVEL | NOTATION |
|---|---|---|
| I. Linguistic modalities | 1. Static analogue sign graphic language | <li,an,-ar,sta,gra> |
| | 2. Static analogue sign acoustic language Dynamic analogue sign acoustic language | <li,an,-ar,sta/dyn,aco> |
| | 3. Static analogue sign haptic language Dynamic analogue sign haptic language | <li,an,-ar,sta/dyn,hap> |
| <li,-an,-ar> | 4. Dynamic analogue sign graphic language | <li,an,-ar,dyn,gra> |
| | 5. Static non-analogue sign graphic language | <li,-an,-ar,sta,gra> |
| | 6. Static non-analogue sign acoustic language Dynamic non-analogue sign acoustic language | <li,-an,-ar,sta/dyn,aco> |
| | 7. Static non-analogue sign haptic language Dynamic non-analogue sign haptic language | <li,-an,-ar,sta/dyn,hap> |
| | 8. Dynamic non-analogue sign graphic language | <li,-an,-ar,dyn,gra> |
| II. Analogue modalities | 9. Static analogue graphics | <-li,an,-ar,sta,gra> |
| | 10. Static analogue acoustics Dynamic analogue acoustics | <-li,an,-ar,sta/dyn,aco> |
| <-li,an,-ar> | 11. Static analogue haptics Dynamic analogue haptics | <-li,an,-ar,sta/dyn,hap> |
| | 12. Dynamic analogue graphics | <-li,an,-ar,dyn,gra> |
| III. Arbitrary modalities | 13. Arbitrary static graphics | <-li,-an,ar,sta,gra> |
| | 14. Arbitrary static acoustics Dynamic arbitrary acoustics | <-li,-an,ar,sta/dyn,aco> |
| <-li,-an,ar> | 15. Arbitrary static haptics Dynamic arbitrary haptics | <-li,-an,ar,sta/dyn,hap> |
| | 16. Dynamic arbitrary graphics | <-li,-an,ar,dyn,gra> |
| IV. Explicit modality structures | 17. Static graphic structures | <-li,-an,-ar,sta,gra> |
| | 18. Static acoustic structures Dynamic acoustic structures | <-li,-an,-ar,sta/dyn,aco> |
| | 19. Static haptic structures Dynamic haptic structures | <-li,-an,-ar,sta/dyn,hap> |
| <-li,-an,-ar> | 20. Dynamic graphic structures | <-li,-an,-ar,dyn,gra> |
| SUPER LEVEL | GENERIC UNIMODAL LEVEL | NOTATION |

**Table 1. Bernsen's ontology taxonomy [79], "li": linguistic; "an": analogue; "ar": arbitrary; "sta": static; "dyn": dynamic; "gra": graphics; "aco": acoustics; "-": not**

Bernsen's modality taxonomy describes the representational features of modalities, which determine their capabilities to represent types of information. However, it does not address the perceptual features of modalities, which determine the way each modality is perceived and processed by human cognitive systems. Bachvarova [84] argued that the modality classification should be extended to 3 levels:
  1) The information presentation level models the properties of modalities which determine their capabilities to represent different types of information

2) The perception level models the properties that determine the way each modality is perceived and processed by human cognitive systems

3) The structure level models the dependencies between different modalities

The first level supports mainly the modality allocation process, while the second and the third level are used for calculating modality combinations. In this research, we combine these two modality taxonomies/ontologies. We use Bernsen's modality taxonomy to describe the available modalities and their perception properties are taken as one of the factors during modality planning process.

## 6.1.3  The modality planning module

The function of the modality planning module is to calculate the optimum modality utilization to achieve an effective and efficient presentation manner (see p.11). The input of this module is a series of action instances generated from the on-going crisis event (see section 5.2). Different modality planning strategies are applied to different types of action. The output of this module contains the modality selection results and the inspections of how to generate instances for each selected modality. The instance-generation inspections also indicate the spatial and temporal  combination manner of the selected modalities. For instance, if two modalities are assigned the same starting time, theyFinally, the outputs of this module are passed on to the text generator, the graphics generator and the speech generator. They take their own relevant parts and implement the presentation.

### *6.1.3.1  The available modalities for planning*

Both visual and auditory modalities are available for planning. Visual modalities include map, text, and image. Auditory modalities include speech and sound effects. Sutcliffe et al. [85] have introduced a set of attention effect advices for directing the user's attention to the appropriate information at the correct level of detail. Following these advices, dynamic text and dynamic image are used when extra attention is needed. Based on Bernsen's modality taxonomy [79], the properties of the available modalities are listed in table 2.

| Modality | Properties |
|---|---|
| Static Text | (li,-an,-ar,sta,gra) |
| Dynamic Text | (li,-an,-ar,dyn,gra) |
| Map | (-li,an,-ar,sta,gra) |
| Static Image | (-li,an,-ar,sta,gra) |
| Dynamic Image | (-li,an,-ar,dyn,gra) |
| Speech | (li,-an,-ar,dyn,aco) |
| Sound Effect | (-li,-an,ar,sta/dyn,aco |

Table 2: The properties of the available modalities (based on [79])

In order to specify the detailed utilization of the modalities, modality models are constructed for text, map, image, speech, and alarm sound, respectively. The modality model contains a set of parameters which describes the utilization details of the modality (see table 3). It can be viewed as a template for creating modality instances. Here, we don't separate static use and dynamic use. These properties are described by the

parameter value. Therefore, static text and dynamic text share the same modality model. The same goes for static images and dynamic images.

| Unimodality | Model Parameters |
|---|---|
| Text | Content, ReferTo, Style, Size, Color, Blink, StartTime, Duration, DisplayArea, ScrollDirection, ScrollSpeed |
| Map | Country, Province, City, InvolvedArea, DisplayedArea |
| Image | Source, ReferTo, DisplayArea, StartTime, Duration, Blink |
| Speech | Content, ReferTo, Tone, Speed, StartTime, RepeatTime |
| Sound Effect | Source, ReferTo, StartTime, RepeatTime |

**Table 3: The modality models**

When fulfilling a specific presentation task, one or more modality instances will be created. Their parameter values also indicate the combination manner. For example, the presentation task is to show the location of the policeman. The modality planner locates an image of a policeman on the map together with a text explanation. The image and text instances are created as follows. The parameter values are filled by the selected modality planning strategy (see section 6.1.3.2). The values of "StartTime", "DisplayArea", and "ReferTo" parameters indicate that the two modality instances will be shown at the same time, near to each other, and both refer to the policeman.

Text-1
- Content:       In Gate Street, 550M to tunnel
- ReferTo:       Policeman.Location
- Style:         Arial, bold
- Size:          Middle
- Color:         Black
- Blink:         N
- StartTime:     immediate, align with Image-1
- Duration:      30 seconds
- DisplayArea: Rectangle [DisplayCoordination(300,560), DisplayCoordination(450,590)]
- ScrollDirection: N
- ScrollSpeed:     N

Image-1
- Source:        policeman.jpg
- ReferTo:       policeman
- DisplayArea: Rectangle [DisplayCoordination(350,500), DisplayCoordination(400,550)]
- StartTime:     immediate, align with Text-1
- Duration:      30 seconds
- Blink:         Y


### 6.1.3.2   The modality planning strategies

The design of the modality planning strategies aims at achieving the presentation goal, i.e. effectiveness and efficiency. The desired presentation manner conveys the information content correctly and helps to prevent cognitive overload. A design proposal is presented in this section. Each presentation task has its own modality planning strategy. A strategy contains three items: 1) a suitable modality list, 2) a default strategy, and 3) a light strategy. The choice between the default strategy and the light strategy is based on the

user's cognitive state. When the user's preference is available, an adapted version of the default strategy will be generated.

**The suitable modality** list indicates which modalities are suitable for contributing to a certain type of presentation task and what each suitable modality expresses. The values of the "ReferTo" parameter will be filled in. In our crisis management application, the map is always shown as background on the display. However, it will be listed as a suitable modality only when the presentation of an action type needs to make use of it. Recall the example of showing the location of the policeman. Sound effects can do little to show a location. Image and text are selected as suitable modalities. The image refers to the policeman and the text refers to the location of the policeman.

**The default strategy** and its adaption The default strategy is designed to achieve the optimum presentation manner for a certain type of task. First, one or more suitable modalities will be selected. Based on the dual-coding theory [81], if the suitable modality list contains both visual modalities and auditory modalities, their combination owns higher priority. Second, the default strategy contains a specification of how to generate modality instances of all the selected modalities. As mentioned before, the parameters of these modality instances also indicate their combination manner. Third, following the attention effect advices in [85], this strategy also attempts to attract the user's attention to what is being presented. For instance, fire alarm (sound effect) is used for a fire report; ambulance alarm is used for a victim report. When necessary, the speech speed is increased with a warning tone. Dynamic texts and dynamic images are also often used. The default strategy will be applied when the user has no specific preference and no cognitive overload of the user is recognized. If the user especially prefers certain modalities, an adapted version of the default strategy will be made. The adaption to cognitive overload will be described in the following subsection. The adaptation to a user's preferences intends to avoid undesired annoyance for the user. If the user prefers a certain modality, it will always be selected, as long as it is on the list of suitable modalities. The user can also indicate that he prefers a less intrusive presentation manner. Then attraction efforts (e.g. using sound effect, warning tone etc.) will be reduced. The user's preferences are set up before using the system, but not during the crisis management process.

**The Light Strategy** is less attracting. It often contains only visual modalities and it will be applied when the system recognizes that the user might be experiencing cognitive overload. When cognitive overload occurs, the user might become slow at responding to the newly-presented information. The system will notice that more requests/reports stay in pending state. For instance, a victim report stays in the pending state until the system hears a command to the doctor, addressing this victim. When possible cognitive overload is detected, the system still continues on presenting new-coming information. However, only a few most urgent tasks are presented with the default strategy, the rest will adopt the light strategy. In this way, the user's attention is drawn to only the most urgent issues. When a light-presented task becomes one of the most urgent tasks, its presentation will be refreshed with the default strategy. When the user's cognitive state recovers, all light-presented tasks will be shown using the default strategy.

## 6.2  Order planning

The assumption of the order planning module is that the order in which the information is presented has influence on the user's understanding performance (correctness, easiness, time consumption etc.). First, the user may start to analyze the information and build his understanding as soon as the presentation starts. Once an understanding is forming, there is a chance that he will have the tendency to stick to it by biased information selection. Therefore, the information which is presented in the early stage of a presentation might be more critical to the user's situation diagnose. Second, there are several orders a presentation can follow, such as the time sequence, the causal chain, and the sub-story line following a certain object or actor. Then, the question arises is that, at certain point of a presentation, how to decide which data instance should be presented next. Should it be the one with the earliest time stamp? Should it be the one which is the direct consequence of the data instance just presented? Is it better for understanding if the presentation continues on the same story line for some time period and then goes backwards to another story line? At the moment of writing, we don't have clear answers for the above-mentioned questions. However, an intuitive answer would be that a presentation order which assists a correct and coherent understanding should be a good tradeoff among the three possible orders.

Our assumption of the order effect has been proved by an experimental study from Schraagen et al. [32]. They designed a critical thinking tool to support decision making in crisis situations. A user experiment has been carried out to compare the crisis thinking tool with no support conditions. Users acted as crisis managers in the experiment. Their task was to interpret incoming information of an explosion crisis scenario and determine the likely cause of the explosion. The order effect was evaluated in two levels: 1) no framing condition: messages appeared in random, though fixed, order; 2) framing condition: messages supportive of a particular hypothesis appeared in one block after each other. Results show a large and significant performance improvement in the framing condition. Users under framing conditions developed more reliable stories in their minds, which in turn enabled more frequent confirmations of the correct information and less frequent confirmations of non-informative messages.  This result strongly confirmed that the presentation order is one of the important factors for the prevention of tunnel vision. However, it is insufficient to answer the above-mentioned questions. Research effort will be further devoted to this issue.


# 7  Context reasoning and graph-based data mining

As mentioned in chapter 2.2, crisis managers typically have to deal with information overload and large uncertainty. Information often contains redundancies, errors and contradictions. Inconsistent information chunks need to be pieced together into a coherent understandable picture of the reality. Aiming at assisting this process, the reasoning and mining module of the system intends to enhance the information quality and reduce the quantity by applying context reasoning technologies and graph mining technologies. The

functions of this module include 1) data cleaning in order to reduce the information quantity and deal with contradictions and errors, 2) context reasoning in order to enhance the consistency of the data, using semantic-based inference rules and domain specific rules, and 3) graph mining in order to extract graph contents or sub-graphs which are relevant to a user query. Figure 5 shows the structure of the envisaged reasoning and mining module. The four main components are the data cleaner, the context reasoner, the state-change generator and the graph miner.

The functions of these four components will be explained in more detail in the remainder of this section. However, a more elaborated design proposal is out of the scope of the author's research focus.
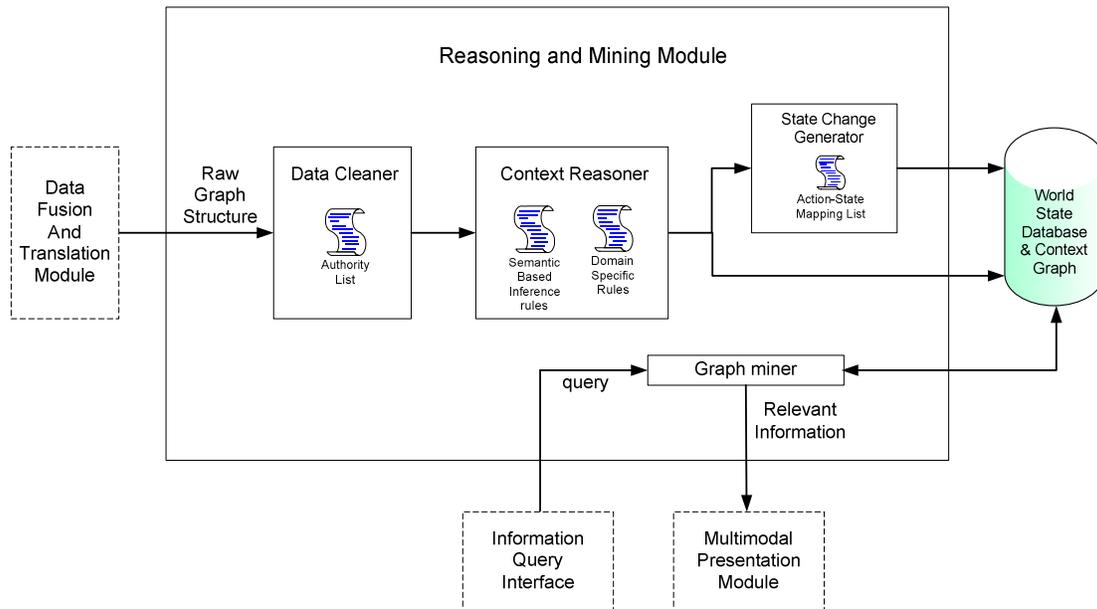


**Figure 5: Reasoning and Mining Module Structure**

## 7.1  Data cleaner

The data cleaner deals with the redundancy, error and contradictions in the raw data instances. If certain information is repeated more than once, all data instances containing this piece of information will be grouped into one. The data cleaner has no general understanding of the real crisis situation; therefore it is not its function to make judgment on contradictions. However, it provides statistical calculation based on an authority list which indicates the experts associated with a certain type of object. For instance, the fireman's expertise is linked with the fire object and the doctor's expertise is linked with the wounded people. According to the list, the data cleaner combines the contradictory data instances into one with a probabilistic trust weight. Information coming from experts is given higher trust weight. In the example shown in Table 4, a fire happened on the 5th floor of the building A in the street B. The control room first received fire reports from two civilians in the street and one civilian in the building, but there were contradictions on the location of the fire. The reason might be that the heavy smoke made it difficult for

26

people in the street to locate the fire. The data cleaner combined the three fire reports as one, giving the location property two values with different trust weights. After a while, doctor and fireman both arrived and reported the fire status. There was again contradiction on the location of the fire, but the data cleaner trusted the report from the fireman who was not only an expert but also inside the building. Therefore, the doctor's report was discarded.

| Time slot | time stamp t1 ~ t2 | | | time stamp t3~t4 | |
|---|---|---|---|---|---|
| **Actor** | civilian1<br>- street B | civilian2<br>- street B | civilian3<br>-   building A | doctor<br>-   street B | fireman<br>- building A |
| **Raw data instance** | Report<br>- fire<br>- building A<br>- floor 4 | Report<br>- fire<br>- building A<br>- floor 5 | Report<br>- fire<br>- building A<br>- floor 5 | Report<br>- fire<br>- building A<br>- floor 4 | Report<br>- fire<br>- building A<br>- floor 5 |
| **Data cleaner output** | Report<br>- fire1<br>- building A<br>- floor 5(70%)<br>  floor 4(30%) | | | Report<br>- fire1<br>- building A<br>- floor 5 | |

**Table 4: An example of data cleaning**

## 7.2  Context reasoner

The context reasoner reasons about the context information the data contains, based on semantic-based inference rules and domain specific rules. The goal is to enhance the data consistency. Semantic-based inference rules are often entailed by the knowledge representation framework/language. The domain specific rules are the general practice knowledge in a specific application domain. For instance, in the U-HealthCare information system presented by Ko et al. [86], the knowledge representation is based on the Web Ontology Language (OWL). Therefore, the OWL axiom is used as the semantic-based infer rules. Doman specific rules are the healthcare service rules. They are used to inference personalized health care strategies from the physiological status of a patient. Wang et al. [87] present a context model in the home domain. They use the DL (description logic) reasoning rules entailed by OWL to reason about the physical locations of people at home. They also defined a set of home domain specific rules to enable the reasoning for questions like "what the user is doing?"

Suppose that we use OWL to implement our world ontology and scenario representation structure. Then, DL reasoning rules can be applied to reason about the consistency of classes and the properties of data instances. Table 5 shows a selection of DL reasoning rules for instance property reasoning. "P (A, B)" stands for "property P of instance A has value B". Notice that, in OWL, a data instance can have another data instance as property value.

| Transitive property P | $P(A, B) \wedge P(B, C) \vdash P(A, C)$ |
|---|---|
| Functional property P | $P(A, B) \wedge P(A, C) \vdash B=C$ |
| Inverse Functional property P | $P(A, B) \wedge P(C, B) \vdash A=C$ |
| Symmetric property P | $P(A, B) \vdash P(B, A)$ |

**Table 5: A selection of DL reasoning rules for instance property reasoning**

An example is given for further explanation. Suppose a doctor drives towards a victim after he receives a rescue command. At a certain moment, the automatic location report from the doctor's PDA matches the location of the victim. As the property "location" is functional, the system fulfills the following inference and concludes that the doctor has reached the victim.

$$\text{location ( doctor, location A ) } \Lambda \text{ location ( victim, location A )}$$
$$\vdash \text{doctor has reached victim}$$

During the crisis management training, the crisis managers are educated by a set of general guidelines associated with a certain type of crisis. They are mostly the experiences learned from historical cases. Although the rules might not always hold for each new crisis case, it is still dangerous if the crisis managers don't take them into consideration during their decision making processes. These domain specific rules can be translated into inference rules and applied during the context reasoning. When necessary, warnings will be presented together with the data instance. Take the real Swiss tunnel fire crisis in 2001 as an example. Firemen were sent into the tunnel to rescue the victims. Due to extremely high temperature and heavy smoke, they were not able to locate themselves and communicate with each other. Finally, they lost both the victims and several firemen. Lesson learnt from this typical decision mistake is that the crisis managers should seriously consider the temperature and the visibility condition before they send the firemen into the fire. This rule can be translated into the following rule first-order logic rule.

$$\text{Temperature ( location B, >100 °C ) } \Lambda \text{ Visibility ( location B, < 1m )}$$
$$\vdash \text{Warning ( location B, dangerous for entering)}$$

## 7.3  State change generator

The state change generator creates the instances of *state_change* class according to an action-state mapping list and modifies the world state database. Table 6 shows a sample of the action-state mapping list. The data format of *action* and *state_change* instance look pretty much the same, however, they need to be distinguished for two reasons. First, some action types don't have any direct influence on the world state, such as commands, requests and answers; therefore there are no *state_change* instances related to them. Second, *action* instances are processed by the data cleaner before they are mapped into a *state_change* instance.

## 7.4  Graph miner

As the crisis event develops, the context graph is extended by new-coming data instances. Via a query interface, the user is able to access the previous information. A user query consists of several keywords referring to a certain entity/action and the time stamp/interval. The function of the graph miner is to extract relevant contents from the large scale structured database based on a user query.

Graph-based data mining is a new branch of data mining. Though it is a pretty young research area, it has gained a great deal of research effort from many different areas, such as semantic web [88-90], bioinformatics [91] and computational chemistry [92, 93]. Conventional data mining algorithms mainly deal with the discovery of patterns in collections of Independently and Identically Distributed (IID) instances. However, real-life large scale databases, such as molecule databases, email databases and webpage databases, often contain a variety of object types and rich link structure. They are described as networks and graphs. Graph-based data mining not only deals with the patterns in the data instances, but also mines the information embedded in the topological structures.

Graph mining in this research can use either an object-based method, or a link-based method, or a combination of these two. A simple and intuitive object-based method would be to select all data instances which match the required key words. However, a large database requires an efficient graph search method in order to guarantee the speed of the searching process. Addressing this problem, there are five groups of graph searching approaches [94]: the greedy search based approach, the inductive logic programming (ILP) based approach, the inductive database based approach, the mathematical graph theory based approach, and the kernel function based approach. Lind-based mining is to exploit the link structure of a graph in order to prioritize a set of objects within the graph [95]. The most notable application is the web page ranking in the context of web information retrieval. Two basic algorithms which might be also plausible for this research are PageRank [96] and HITS [97]. A method combining the object-based method and the link-based method might be able to enhance the coherence of the searching results.

| | | | |
|---|---|---|---|
| **Action** | class: Find<br>Instance: f<br>- time<br>- finder<br>- object<br>- property1<br>- property2<br>- … | Class: Report<br>Instance: r<br>- time<br>- reporter<br>- object<br>- property<br>- new value | Class: Solve<br>Instance: s<br>- time<br>- actor<br>- object |
| **State Change** | Class: New Object<br>Instance: object(f)<br>- name: object(f)<br>- found at: time(f)<br>- found by: founder(f)<br>- property1: property1(f)<br>- property2: property2(f)<br>- … | Class: Change<br>Instance: object(r)<br>- time: time(r)<br>- reported by: reporter(r)<br>- property: property(r)<br>- value: new value(r) | Class: Change<br>Instance: object(s)<br>- time: time(s)<br>- solved by: actor(s)<br>- urgency level *: -1 |

**\*** Negative urgency level indicates that the object is already solved. For instance, fire has been already put off; victim has been rescued, etc.

**Table 6: A sample of action-state mapping list**

# 8  Conclusion

During crisis management process, crisis managers (decision makers) typically have to deal with high urgency, high risk, high uncertainty and high information density. These challenges often lead to cognitive biases in their situation understanding and information confirmation processes, so called "tunnel vision". In this report, we proposed an information assistant system to prevent tunnel vision. The information assistant system serves as a platform for monitoring the on-going crisis event. All data goes through the system before it arrives at the user. The *multimodal data fusion module* interprets the combined semantics of multimodal inputs and reduces the noises and errors a single input modality may contain. The *reasoning and mining module* applies data cleaning and context reasoning to the raw data instances, in order to reduce the information quantity and enhance the consistency. The development of the crisis scenario is recorded in the *system memory*. Via an *information query interface*, users are able to access the previous facts. In this way, the system memory serves as an extension of the human working memory. The *multimodal presentation generation module* intends to convey the crisis information in an effective and efficient manner. Effectiveness means that the information is conveyed correctly; while efficiency means that the presentation manner helps to prevent or repair the user's cognitive overload. Two main components are the *order planning module* and the *modality planning module*. The order planning module determines a presentation order which enables a correct and coherent understanding. The modality planning module calculates the optimum utilization of the available modalities. The presentation manner adapts to the user's preferences and cognitive state. In summary, the system enhances the data quality, reduces the data quantity and presents the crisis information in a manner that prevents or repairs the user's cognitive overload. Working with such an information assistant system, the human managers are more likely to keep a clear mind about the actual situation and stay open-mined on their decisions.

# 9  Acknowledgement

# References

[1]     K. Fearn-Banks, *Crisis Communications : A Casebook Approach*: Lawrence Erlbaum Association, 1996.

[2]     C. M. Pearson and J. A. Clair, "Reframing Crisis Management," *The Academy of Management Review,* vol. 23, pp. 59-76, 1998.

[3]     K. Smith, *Environmental Hazards: Assessing Risk and Reducing Disaster*: Routledge, 2001.

[4]     C. E. Fritz, *Disaster*: Institute for Defense Analyses, Weapons Systems Evaluation Division, 1961.

[5]     T. E. Drabek and G. J. Hoetmer, *Emergency Management: Principles and Practice for Local Government*: International City Management Association, 1991.

[6]     L. J. Dunn, "HCI factors affecting quality of information in crisis management systems," in *Sixth Australian Conference on Computer-Human Interaction*, 1996, pp. 324-325.

[7]     J. S. Horsley and R. T. Barker, "Toward a Synthesis Model for Crisis Communication in the Public Sector: An Initial Investigation," *Journal of Business and Technical Communication,* vol. 16, p. 406, 2002.

[8]     L. Carver and A. M. Turoff, "Human-computer interaction: the human and computer as a team in emergency management information systems " *Commun. ACM* vol. 50, pp. 33-38, 2007.

[9]     J. F. Nunamaker Jr, E. S. Weber, C. A. P. Smith, and M. Chen, "Crisis planning systems: tools for intelligent action," *System Sciences, 1988. Vol. III. Decision Support and Knowledge Based Systems Track, Proceedings of the Twenty-First Annual Hawaii International Conference on,* vol. 3, 1988.

[10]    S. Belardo and H. L. Pazer, "A framework for analyzing the information monitoring and decision support system investment tradeoff dilemma: an application to crisis management," *Engineering Management (IEEE Transactions),* vol. 42, pp. 352-359, Nov. 1995.

[11]    M. Turoff, M. Chumer, B. Van de Walle, and X. Yao, "The Design of a Dynamic Emergency Response Management Information System (DERMIS)," *Journal of Information Technology Theory and Application,* vol. 5, pp. 1-35, 2004.

[12]    B. M. Staw, L. E. Sandelands, and J. E. Dutton, "Threat-Rigidity Effects in Organizational Behavior: A Multilevel Analysis," *Administrative science quarterly,* vol. 26, p. 501, 1981.

[13]    R. Desimone, "The application of AI planning technology to crisis management tasks," in *Intelligent Planning and Scheduling Solutions, IEE Colloquium on*, 1996, pp. 6/1-6/3.

[14]    N. R. Jennings and M. J. Wooldridge, *Agent technology: foundations, applications, and markets*: Springer-Verlag New York, Inc. Secaucus, NJ, USA, 1998.

[15]    F. Fiedrich and P. Burghardt, "Agent-based systems for disaster management," *Commun. ACM,* vol. 50, pp. 41-42, 2007.

[16]    Y. Murakami, K. Minami, T. Kawasoe, and T. Ishida, "Multi-agent simulation for crisis management," *Knowledge Media Networking, 2002. Proceedings. IEEE Workshop on,* pp. 135-139, 2002.

[17]    G. Santos and B. E. Aguirre, "A Critical Review of Emergency Evacuation Simulation Models," *Proceedings of Building Occupant Movement during Fire Emergencies,* pp. 10-11, 2004.

[18]    A. Smirnov, M. Pashkin, T. Levashova, N. Shilov, and A. Kashevnik, *Role-Based Decision Mining for Multiagent Emergency Response Management* vol. 4476/2007. Berlin / Heidelberg: Springer, 2007.

[19] F. Fiedrich, "An HLA-based multiagent system for optimized resource allocation after strong earthquakes," *Proceedings of the 37th conference on Winter simulation,* pp. 486-492, 2006.

[20] N. R. Jennings, S. D. Ramchurn, P. Dutta, A. Rogers, and I. Vetsikas, "The ALADDIN project: Agent technology to the rescue," *Proceedings of the 1st ATDM Workshop.*

[21] B. Van Veelen, P. Storms, and C. van Aart, "Effective and efficient coordination strategies for agile crisis response organizations," *Proceedings of ISCRAM 2006.*

[22] H. Kitano and S. Tadokoro, "RoboCup rescue: A grand challenge for multiagent and intelligent systems," *AI MAG,* vol. 22, pp. 39-52, 2001.

[23] S. Chapman and F. Ciravegna, "Focused Data Mining for decision support in Emergency Response Scenarios," *Proc. Int. Semantic Web Conference 2006 Workshop on Web Content Mining with Human Language Technologies,* 2006.

[24] N. Krahnstoever, E. Schapira, S. Kettebekov, and R. Sharma, "Multimodal human-computer interaction for crisis management systems," *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on,* pp. 203-207, 2002.

[25] R. Sharma, M. Yeasin, N. Krahnstoever, I. Rauschert, G. Cai, I. Brewer, A. M. MacEachren, and K. Sengupta, "Speech-gesture driven multimodal interfaces for crisis management," *Proceedings of the IEEE,* vol. 91, pp. 1327-1354, 2003.

[26] S. Fitrianie, R. Poppe, T. H. Bui, A. G. Chitu, D. Datcu, D. H. W. Hofs, D. J. M. Willems, M. Poel, L. J. M. Rothkrantz, L. G. Vuurpijl, and J. Zwiers, "Multimodal Human-Computer Interaction in Crisis Environments," in *the 4th International ISCRAM Conference*, Delft, the Netherlands, 2007.

[27] S. Kettebekov, N. Krahnstoever, M. Leas, E. Polat, H. Raju, E. Schapira, and R. Sharma, "i2Map: Crisis Management using a Multimodal Interface," *ARL Federate Laboratory 4th Annual Symposium,* 2000.

[28] R. N. Hasan, Samadzadeh, M.H. , " A study of tunnel vision and rotation as aspects of Web site visibility," in *International Conference on Information Technology: Coding and Computing.( ITCC 2005)* 2005, pp. 509-514.

[29] D. A. Norman, "Emotion & Design: Attractive things work better," in *Interactions Magazine*, 2002, pp. 36-42.

[30] L. M. Dianne, "Lessons About Justice from the "Laboratory" of Wrongful Convictions: Tunnel Vision, the Construction of Guilt and Informer Evidence," *70 UMKC L. REV. 847,848,* 2002.

[31] A. F. Keith, Scott, M. , "The Multiple Dimensions of Tunnel Vision in Criminal Cases," *Wisconsin Law Review,* vol. 2, 2006.

[32] J. Schraagen, A. Eikelboom, and G. te Brake, "Experimental evaluation of a critical thinking tool to support decision making in crisis situations," *Proceedings of the 2nd International Conference on Information Systems for Crisis Response and Management, Brussels, Belgium, 18¨C20 April 2005,* 2005.

[33] G. A. Klein, *Sources of Power: How People Make Decisions*: Cambridge, MA:MIT Press, 1998.

[34] N. Pennington and R. Hastie, "The story model for juror decision making," *Inside the juror: The Psycholigy of Juror Decision Making,* pp. 192-221, 1993.

[35]  W. H. Starbuck, A. Greve, and B. L. T. Hedberg, "Responding to Crisis," *Journal of Business Administration,* vol. 9, pp. 111-37, 1978.

[36]  R. E. Rice, "Computer-mediated communication and organizational innovation," *Journal of Communication,* vol. 37, pp. 65-94, 1987.

[37]  A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science,* vol. 185, p. 1124, 1974.

[38]  J. Sweller, "Cognitive Load During Problem Solving: Effects on Learning," *Cognitive Science,* vol. 12, pp. 257-285, 1988.

[39]  P. Chandler and J. Sweller, "Cognitive Load Theory and the Format of Instruction," *Cognition and Instruction,* vol. 8, pp. 293-332, 1991.

[40]  J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learning and Instruction,* vol. 4, pp. 295-312, 1994.

[41]  F. Paas, A. Renkl, and J. Sweller, "Cognitive Load Theory and Instructional Design: Recent Developments," *Educational Psychologist,* vol. 38, pp. 1-4, 2003.

[42]  A. A. Esgate, *An introduction to applied cognitive psychology*, 2005.

[43]  R. Azuma, Daily, M., Furmanski, C., "A review of time critical decision making models and human cognitive processes," in *Aerospace Conference, IEEE*, 2006, p. 9 pp.

[44]  P. Papaioannou and G. Georgiou, "Human behaviour in tunnel accidents and incidents: endusers, operators and response teams " 2003.

[45]  H. J. Einhorn and R. M. Hogarth, "Behavioral Decision Theory: Processes of Judgment and Choice," *Annual Review of Psychology,* vol. 32, pp. 53-88, 1981.

[46]  M. M. Mowen and J. C. Mowen, "An empirical examination of the biasing effects of framing on business decisions," *Decision Sciences,* vol. 17, pp. 596-602, 1986.

[47]  A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science,* vol. 211, pp. 453-458, 1981.

[48]  G. Wright and P. Goodwin, "Eliminating a framing bias by using simple instructions to 'think harder' and respondents with managerial experience: Comment on 'Breaking the frame'," *Strategic management journal,* vol. 23, pp. 1059-1067, 2002.

[49]  J. Evans, *Bias in Human Reasoning: Causes and Consequences*: Erlbaum, 1989.

[50]  T. C. Chamberlin, "The Method of Multiple Working Hypotheses," *Science,* pp. 754-759, 1965.

[51]  T. R. Gruber, "Towards Principles for the Design of Ontologies Used for Knowledge Sharing," *Int. Journal of Human-Computer Studies,* vol. 43, pp. 907--928, 1995.

[52]  N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Knowledge Systems Laboratory, Stanford University 2001.

[53]  S. Fitrianie and L. J. M. Rothkrantz, "Technical Report of ICIS Project: Knowledge representation of the world model," 2006.

[54]  I. Gurevych, R. Porzel, E. Slinko, N. Pfleger, J. Alexandersson, and S. Merten, "Less is more: using a single knowledge representation in dialogue systems," in *HLT-NAACL 2003 workshop on Text meaning*, 2003, pp. 14--21.

[55]  A. K. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing,* vol. 5, pp. 4-7, 2001.

[56]   P. Brézillon, "Focusing on Context in Human-Centered Computing," *IEEE Intelligent Systems,* vol. 18, pp. 62-66, 2003.

[57]   T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, "An Ontology-based Context Model in Intelligent Environments," *Proceedings of Communication Networks and Distributed Systems Modeling and Simulation Conference,* vol. 2004, 2004.

[58]   Y. Li, J. I. Hong, and J. A. Landay, "contextMap: modelling Scenes of the real world for context-aware computing," *http://guir.berkeley.edu/pubs/ubicomp2003/ contextmap.pdf*.

[59]   P. Brézillon, "Context-based modeling of operators' practices by contextual graphs," *Proceedings of the 14th Mini Euro Conference: Human Centered Processes,* pp. 129–137, 2003.

[60]   G. K. Mostefaoui and P. Brezillon, "Modeling context-based security policies with contextual graphs," *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on,* pp. 28-32, 2004.

[61]   Y. Bachvarova, L. Hardman, and J. V. Ossenbrunggen, "Modeling the process of selecting modalities in automatic generation of multimedia presentations." *First International Workshop on Interaction Design and the Semantic Web (IDSW 2004)*, 2004

[62]   R. Sharma, V. I. Pavlovic, and T. S. Huang, "Toward multimodal human-computer interface," *Proceedings of the IEEE,* vol. 86, pp. 853-869, 1998.

[63]   S. L. Oviatt and E. Olsen, "Integration Themes in Multimodal Human-Computer Interaction," *Proceeding of the International Conference on Spoken Language Processing,* vol. 2, pp. 551-554, 1994.

[64]   S. Oviatt, "Multimodal interfaces for dynamic interactive maps," *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground,* pp. 95-102, 1996.

[65]   N. Elouazizi and Y. Bachvarova, "On Cognitive Relevance in Automatic Multimodal Systems," *Multimedia Software Engineering, 2004. Proceedings. IEEE Sixth International Symposium on,* pp. 418-426, 2004.

[66]   C. Karagiannidis, A. Koumpis, and C. Stephanidis, "Media/Modalities Allocation in Intelligent Multimedia User Interfaces: Towards a Theory of Media and Modalities," *First International Workshop on Intelligence and Multimodality in Multimedia Interfaces (IMMI-1),* 1995.

[67]   Y. Arens, E. Hovy, and S. van Mulken, "Structure and Rules in Automated Multimedia Presentation Planning," *Proc. IJCAI,* vol. 93, pp. 1253-1259, 1993.

[68]   Y. Arens and E. Hovy, "How to Describe What? Towards a Theory of Modality Utilization," *The Twelfth Annual Conference of the Cognitive Science Society,* pp. 487–494, 1990.

[69]   S. Kerpedjiev, G. Carenini, S. F. Roth, and J. D. Moore, "Integrating planning and task-based design for multimedia presentation," *Proceedings of the 2nd international conference on Intelligent user interfaces,* pp. 145-152, 1997.

[70]   N. Clark, "The language of data: tables and graphs as exposition," *Proceedings ICRISA Tt,* 1986.

[71]   R. A. Gargan Jr, J. W. Sullivan, and S. W. Tyler, "Multimodal response planning: an adaptive rule based approach," *Proceedings of the SIGCHI conference on Human factors in computing systems,* pp. 229-234, 1988.

[72]    J. Mackinlay, "Automating the Design of Graphical Presentations of Relational Information," *ACM Transactions on Graphics,* vol. 5, 1986.

[73]    S. K. Feiner and K. R. McKeown, "Automating the Generation of Coordinated Multimedia Explanations," *IEEE Computer,* vol. 24, pp. 33-41, 1991.

[74]    W. Wahlster, E. Andre, W. Finkler, H. J. Profitlich, and T. Rist, "Plan-Based Integration of Natural Language and Graphics Generation," *Artificial Intelligence,* vol. 63, pp. 387-427, 1993.

[75]    W. Wahlster, N. Reithinger, and A. Blocher, "SmartKom: Multimodal Communication with a Life-Like Character," *Seventh European Conference on Speech Communication and Technology,* 2001.

[76]    C. Elting, J. Zwickel, and R. Malaka, "What are Multimodalities made of? Modeling Output in a Multimodal Dialogue System," *International Conference on Intelligent User Interfaces IUI,* pp. 13-16, 2002.

[77]    M. X. Zhou, Z. Wen, and V. Aggarwal, "A graph-matching approach to dynamic media allocation in intelligent multimedia interfaces," *Proceedings of the 10th international conference on Intelligent user interfaces,* pp. 114-121, 2005.

[78]    S. Gold and A. Rangarajan, "A graduated assignment algorithm for graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 18, pp. 377-388, 1996.

[79]    N. O. Bernsen, "Multimodality in language and speech systems – from theory to design support tool," *Multimodality in Language and Speech Systems,* 2002.

[80]    N. O. Bernsen, "Why are Analogue Graphics and Natural Language both Needed in HCI," *Interactive Systems: Design, Specification, and Verification. Focus on Computer Graphics. Springer Verlag,* pp. 235-51, 1995.

[81]    A. Paivio, *Mental Representations: A Dual Coding Approach*: Oxford University Press, USA, 1986.

[82]    S. Y. Mousavi, R. Low, and J. Sweller, "Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes," *Journal of Educational Psychology,* vol. 87, pp. 319-34, 1995.

[83]    N. O. Bernsen, "Foundations of multimodal representations: a taxonomy of representational modalities," *Interacting with Computers,* vol. 6, pp. 347-371, 1994.

[84]    Y. Bachvarova, B. van Dijk, and A. Nijholt, "Towards a Unified Knowledge-Based Approach to Modality Choice," *MOG 2007 Workshop on Multimodal Output Generation*, 2007

[85]    A. G. Sutcliffe, S. Kurniawan, and J. E. Shin, "A method and advisor tool for multimedia user interface design," *International Journal of Human-Computer Studies,* vol. 64, pp. 375-392, 2006.

[86]    E. J. Ko, H. J. Lee, and J. W. Lee, "Ontology-Based Context Modeling and Reasoning for U-HealthCare," *IEICE TRANSACTIONS on Information and Systems,* vol. 90, pp. 1262-1270, 2007.

[87]    X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung, "Ontology based context modeling and reasoning using OWL," *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on,* pp. 18-22, 2004.

[88]    J. Iria and F. Ciravegna, "Relation Extraction for Mining the Semantic Web," *Dagstuhl Seminar on Machine Learning for the Semantic Web,* 2005.

[89]    A. Strehl, J. Ghosh, and R. Mooney, "Impact of similarity measures on web-page clustering," *Proc. AAAI Workshop on AI for Web Search (AAAI 2000), Austin,* pp. 58–64, 2000.

[90]    C. Ramakrishnan, W. H. Milnor, M. Perry, and A. P. Sheth, "Discovering informative connection subgraphs in multi-relational graphs," *ACM SIGKDD Explorations Newsletter,* vol. 7, pp. 56-63, 2005.

[91]    S. Mukherjea and B. Bamba, "BioPatentMiner: An Information Retrieval System for BioMedical Patents," *Thirtieth International Conference on Very Large Data Bases,* pp. 1066-1077, 2004.

[92]    B. Bringmann and A. Karwath, "Frequent SMILES," *Lernen, Wissensentdeckung und Adaptivitat, Workshop GI Fachgruppe Maschinelles Lernen, LWA,* 2004.

[93]    A. Karwath and L. De Raedt, "Predictive Graph Mining," *Discovery Science: 7th International Conference, DS 2004, Padova, Italy, October 2-5, 2004: Proceedings,* 2004.

[94]    T. Washio and H. Motoda, "State of the art of graph-based data mining," *ACM SIGKDD Explorations Newsletter,* vol. 5, pp. 59-68, 2003.

[95]    L. Getoor and C. P. Diehl, "Link mining: a survey," *ACM SIGKDD Explorations Newsletter,* vol. 7, pp. 3-12, 2005.

[96]    L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," Technical report, Stanford Digital Library Technologies Project, 1998, 1998.

[97]    J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM),* vol. 46, pp. 604-632, 1999.

**Appendix A**: Ontology class hierarchy and properties
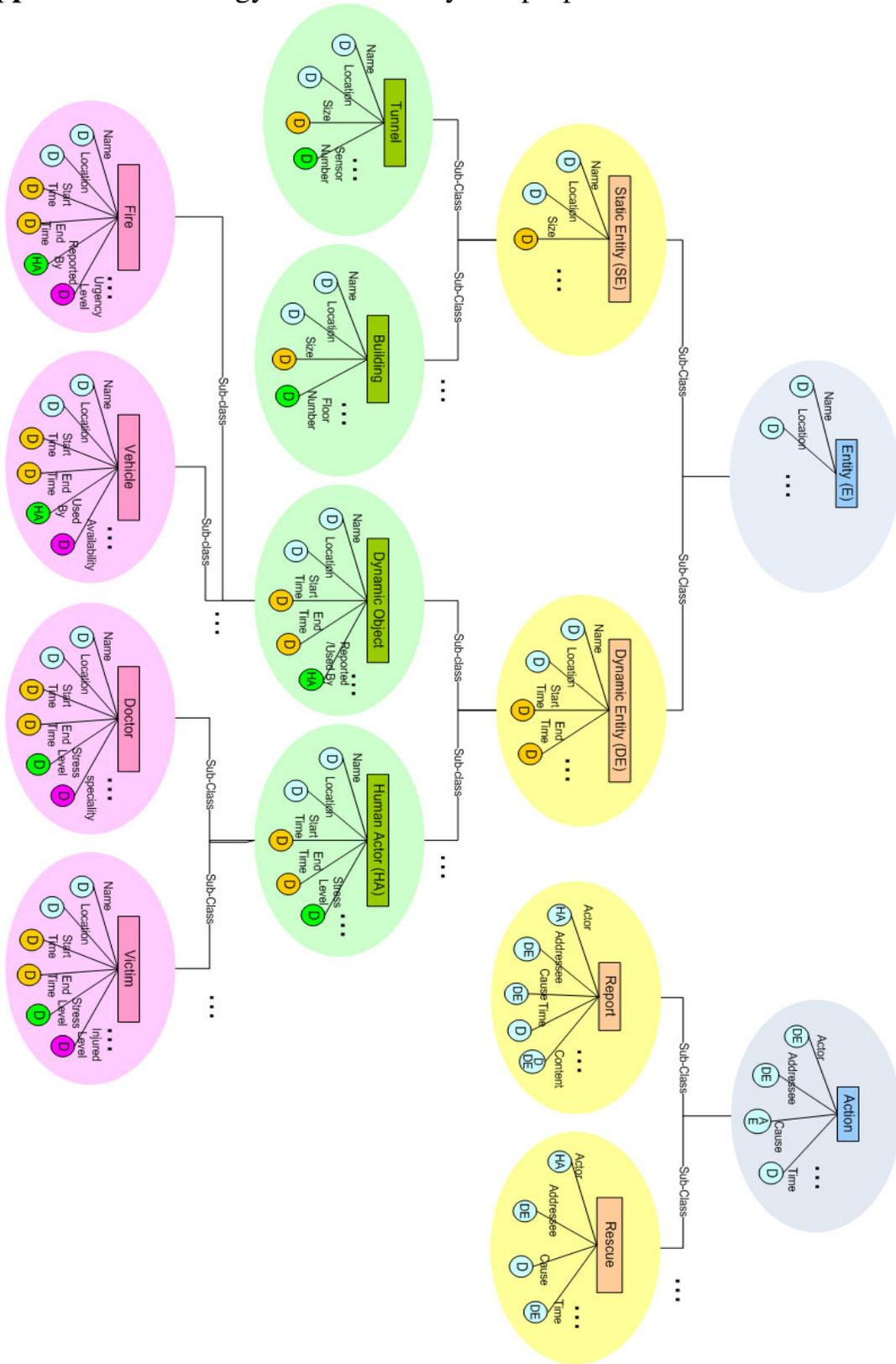
## Appendix B: ICIS common scenario - Tunnel fire

Prologue: a serious accident happened in a huge tunnel. A fireman is already on scene. A doctor is in his car nearby but due to the chaos, he is not able to get anywhere. In a police station not far from the tunnel, a police man is diligently filling out his forms. In the crisis' control room, a controller is available and assists in the communication.

16:45   The accident is detected by sensors and various persons have contacted the alarm number. A broadcast is sent to all cars approaching the tunnel to indicate the accident and make them halt.

16:46   A doctor uses his car's video phone to inform the system of his location and offers his services. He notes that he is stuck in the chaos and needs a police escort to get him to the tunnel. He receives an immediate response from the system that the situation is evaluated and an escort is sought.

16:47   The fireman puts locations of victims on a map and sends these to the system. He receives a request to indicate which victims can be rescued without danger for the rescuer.

16:48   A controller in the crisis' control room is shown a map of the environment of the tunnel with all police stations and the location of the doctor. He is asked to indicate from which police station an escort must be appointed. The controller points at the police station closest to the doctor.

16:49   Within the selected police station, the policeman who looks less stressed is selected by the system and is requested to escort the doctor. On his PDA, an icon map appears with the location of the doctor. A direct communication channel between doctor and police man is set up.

16:50   The fireman has indicated which victims can be rescued without danger. The system communicates the locations of these victims to the police man.

Later:  The policeman escorts the doctor, the doctor saves all lives while the firemen put out all fires. Later that evening they all have a big party and, as it turns out, all victims are even able to make some nice dance moves because they have been rescued so quickly. They all sing: "Chim chim… hooray!"